语音识别技术在广播电视监测中的应用

摘　要：现如今，科学技术飞速发展，广播电视作为传统媒体早已成为人们接收新闻信息的常用途径，在推动广播电视监测监管业务迅速进步方面，最为高效的做法就是将语音识别技术融入到广播电视监测监管之中，从而有效提升广播电视监测监管工作的质量及效率。基于此，为能够确保广播电视监测监管事业的蓬勃发展，理应对广播电视监测监管业务中，语音识别技术的应用和探索给予必要的关注和重视。

关键词：科学技术；语音识别技术；广播电视监测监管；应用

中图分类号：TP241　　**文献标识码：**A

文章编号：1671-0134（2019）01-116-03　　**DOI：**10.19483/j.cnki.11-4653/n.2019.01.032

文 / 赵杨阳

引言

当前，电磁环境越来越复杂，如何有效利用无线电频率资源，最大限度地提升广播电视视频业务的综合质量，充分优化频谱资源的优化配置以更好地为人民群众提供广播电视服务，这些都是广播电视监测监管人员需要负责的重要工作。具体来说，语音识别技术是指借助某种算法，通过计算机来识别语音信号，且把这些信号转换成文本的技术。对于现阶段的语音识别技术来说，广播新闻的识别率为 90% 左右，其中语音识别技术涵盖关键词识别技术、模糊语音技术等多种技术。说话人识别技术也指声纹识别技术，结合语音信号对说话人身份进行识别；关键词识别技术指借助文本信息来获取关键词。在广播电视监测业务中运用语音识别技术，能够对语音信号的特点做出针对性的判断，并智能化地识别频谱占用情况，以此为基础有效监管广播电视的播出内容。

1. 语音识别技术

语音识别技术作为一门综合性交叉型的学科，涵盖了多种领域的学科知识。随着科学技术的进步，语音识别技术得到广泛运用。这项技术利用把输入的声信号转化成其所对应的命令或是文字的一种现代化技术，借助语音识别技术，可以不用传统的键盘、鼠标等操作设备，只需利用语言就能够发号施令，从而更加简便快捷地完成在多种情况下的任务，这项技术在社会中具有十分重要的作用[1]。

自开始研究语音识别技术到现在，语音识别技术的发展已有超过半个世纪的历史。最开始对于语音识别技术的研究只是 Audry 系统的研究，其是那个阶段第一个能够得到几个英文字母的系统。在 20 世纪 60 年代，计算机技术得到发展，促进了语音识别技术进一步发展，利用线性预测分析技术与动态规划技术能够对语音信号出现的模型问题进行有效解决，这一技术是语音识别中最为关键的技术。到了 20 世纪 70 年代，语音识别技术获得重大突破，表现在动态时间规整技术基本成熟，让语音能够等长，同时，矢量量化、隐马尔科夫模型理论得到进一步完善，为语音识别技术的进步奠定了良好的基础。在 20 世纪 80 年代，相关人员更加深入地探究了语音识别技术，提出了多种语音识别算法，最为显著的成就为人工神经网络与 HMM 模型。在 20 世纪 90 年代，语音识别技术得到更加广泛的应用，更多的科技互联网公司均在语音识别技术的研究与开发上投入了庞大的资金。进入 21 世纪后，对于语音识别技术的研究，主要是对多种语种的同声翻译、即兴口语及自然对话进行重点研究[2]。

2. 语音识别方法

语音识别技术常用的方法有：（1）基于语言学与声学的方法。这种方法在语音识别技术中得到最早运用，然而该方法涵盖的知识十分匮乏，这就造成现如今这一方法仍就没有得到大范围应用；（2）随机模型法。现如今，随机模型法已得到比较成熟的应用，这种方法选择提取特征、训练模块、分类模块、判断模块的步骤来识别语音，涉及了隐马尔科夫模型（HMM）理论、动态时间规整（DTW）、矢量量化（VQ）技术，其中最为简便优质的技术为隐马尔科夫模型算法，其在语音识别性能上更加优异，所以大多数语音识别系统都选择隐马尔科夫模型；（3）神经网络的方法。这一方法是在语音识别发展的后期出现的一种新的识别方法，其能够对人类的神经活动进行模拟，并具备人的一些特性，例如自主学习和自动适应等，其归类能力与映射能力比较强，在语音识别技术中具备较高的利用价值，在行业中通过有机结合神经网络与传统的方法，取长补短，从而在极大程度上提升语音识别的效率[3]；（4）概率语法分析法。这种方法属于一种可以识别大长度语段的技术，能够实现区别语言的特征，借助多种层次的知识来解决多种层次的问题，但是这种方法也存在一些不足，最为明显的是需要构建合适、有效的适用知识系统。

3. 语音识别程序

语音识别程序主要包括：（1）语音识别的工作模式：

主要为命令模式与识别模式，结合这两种模式的区别来选择对应类型的程序，从而实现语音识别程序；识别模式的工作原理为引擎系统在后台直接提供一个词库与识别模块库，所有系统无需进一步改动识别语法，只需结合识别引擎所具备的主程序源代码加以改写就能够完成；与之相比，命令模式比较难以实现，必须由程序员来编写词典，再加以编程，然后结合语音词典进行更正与处理；命令模式与识别模式最大的区别是程序员应结合词典内容来修改和核对代码；（2）语音识别环境设置。通常语音识别程序的环境设置主要有：CTI 服务器硬件默认参数采集和设定、识别硬件采集卡初始化、引擎端口设置等，其中应用程序的全部工作均要结合 CTI 技术来开展工作；对于语音采集系统的初始化，语音识别的平台可通过对是否已经输入语音进行判断来开展工作，再利用语音采集系统获得语音，同时利用语音卡来对语音进行输出与采集；在实际工作中，将语音卡中具备的板卡打开，再在程序中加入参数就能够运行；对于引擎端口设置，语音开发平台提供了硬件 API 接口函数，这样做只需要调用与赋值函数就能够运行；（3）编译语音字典。对于语音字典的设置，主要内容有：识别语音的规则、语法、语音模板制作等，结合语音平台的标准来开展。对于语音字典的设置，应先对语音识别核心包进行设置，再结合自己编译的语言标准来实现字典的设置；（4）编制识别主程序。这是语音识别程序编译的最后环节，程序员应为主程序的 Graphical User Interface 界面进行编写，从而方便使用者和计算机开展交互操作[4]。

除此之外，现如今国内外提供语音识别技术的厂商非常多，本文对语音识别平台进行了横向对比，具体见表 1。结合表 1 的内容可以认识到商家的语音识别技术都有一定的优势和特点，企业可结合相应的应用场景来选择恰当的语音识别技术。

4. 广播电视监测监管中语音识别技术的价值

4.1 声纹识别

语音识别技术利用语音波形对人的行为特点进行分析，从而识别出讲话人的身份。一方面，对于讲话人身份的识别，可评估出说话人；另一方面，需要评估判定某一个语音是否为指定人声。语音信号始终是开展声纹识别的基础所在，依据一个人的发音特点对说话人的性格特征进行掌握。作为生物因子认证范畴内的重要内容，声纹识别和我们较为熟悉的指纹有一定的相通之处，都是根据人体生物特征来加以评估的。所不同的是，声纹识别主要是依照说话人的特征加以针对性的识别[5]。

4.2 内容辨识

这里我们所介绍的内容辨识主要是指，以语音的生理属性、物理属性为出发点，结合其内容来进行针对性的分析、鉴别，主要目标是对语音信号承载内容的判断。但是在内容识别方面仍然存在一定的提升空间，如发音习惯、方言等，都对识别效果有着直接的影响；对于这一问题的处理，可利用声纹识别加以解决。为能够使语法、语义和声纹相一致，需要借助词性标记、词语区分、语境理解等综合判定语音内容，不过整体上来说，其需要在短时间内做出大量的对比判断[6]。

4.3 语种与语音的分辨

语种识别指对语音材料兼容的语种特点进行评估，在语种识别的基础上进一步研究而得到语音内容判断与智能翻译技术。同时，语种识别可以评估计算机中多种语音材料，通过识别系统对语音材料加以提取是该技术主要应用的方法。与此同时，标准语音模型和个人语音的对比，是语音评估标准判定过程中明确不标准发音的主要手段。

5. 语音识别技术在广播电视监测的具体应用

5.1 具体应用内容

现阶段，随着语音识别技术的不断突破和发展，自动化语音识别技术的应用，可以帮助我们针对性地判断广播频率节目播出的实时状态，并从中提取关键数据信息，从而对语音类型和语种进行分析，并对言语信号、

表 1　横向比较语音识别技术

语音识别技术	支持平台	在线	离线	收费方式	具体特征
微软 Speech API	Windows	–	离线	免费	Msdn 支撑
Google Speech API	Web	在线	–	免费	仅可用于 Web 在线调用
百度语音	IOS |Android |Web	在线	离线	免费	用于普通话、粤语、四川话
搜狗语音	IOS |Android	–	离线	免费	用于普通话、粤语
微信语音	IOS |Android	在线	–	免费	准确率超过 90%
云知声	IOS |Android |Web |Windows | Linux	在线	离线	收费	音频转写有时间轴功能
讯飞开放平台	IOS | Android | WinPhone | Web | Java |Windows | Linux	在线	离线	免费	用于普通话、粤语及四川话，识别准确率在 95% 以上，全平台 sdk

静音和噪音进行分析，从而实现频道中多个频谱能够同时使用语音识别技术进行充分研究。同时，广播电视监测中引入了电视内容监测与语音识别，这就大大降低了人力资源方面的投入，极大地提升了广播电视监测监管效率[7]。具体表现在以下几个方面：（1）电视监测方面。通过该技术的灵活应用，可以有效构建语音模板、文字模板，进而实现特定语音的识别以及特定文字的识别，并进行实时录像，从而准确判断广播电视播出异态；然后借助监测监管系统将异态上报给监测席位，且发出预警信号，以便工作人员进行及时有效的处理，确保广播电视安全播出；在广播电视监测中运用语音识别技术，可将系统判断率提升至99%，实现广播电视监测智能化；（2）广播监测方面。在广播电视监测中语种识别技术具有十分关键的意义，由于国外电台在我国的播音时间和语种很多，需要花费大量的人力和时间来利用人工对这些语种和播音时间进行实时判断；针对上述问题，我们可以通过卫星实现语音识别库的搜集和接收，然后通过录制音频和语音库中数据的对比，判断出所述语言种类。不过由于语音物理数据差异较大，因此，语言判断应用在实际的使用过程中仍然存在一定的问题。同时，将音频指纹相近度法引入语音识别中来，可以通过大量的学习而适应滤波器，具有信道模型的功能[8]。

5.2 实际案例

主要内容有：（1）信号解调设备。待系统完成监测信号源解调工作之后，将会根据实际情况调用调幅广播解调器、有线电视解调器以及调频广播解调器，从而将所搜集到的广播电视信号解调为普通音频信号，然后通过搜集站对这些音频信号进行录音。具体的解调器和调频广播的数量，可按照监测频道数量进行针对性的选择；（2）信号预处理设备。为最大限度地提升信号的判断效果，可以引入AQC4信号预处理设备，这样就可以通过该设备对音频信号进行预处理，然后通过控制信号处理器对语音信号加以进一步的处理，修正信号失真，为后续操作提供必要的信号源文件，进而完成对频道内容的监测和监管；（3）多路信号搜集站。如上文所述，通过有线电视解调器可解调影像和伴音，从而形成能够为系统所识别的视频信号和模拟音频信号，然后通过信号搜集站对其进行记录，具体的有线电视解调器数量可按照监测频道数量进行恰当安排；除此之外，多路监测信号搜集站可以对广播音频信号和电视信号进行接收，对信号进行搜集和压缩，然后存储到服区前阵列柜中，其中搜集站可分为广播信号和电视信号两种搜集站，前者能够同时进行所有广播信号的采集和记录，并实时呈现音量，配合软调音台不仅能够实现对信号增益水平的有效控制，而且可以实现对所有输入信号的完整监控，定时设计了采集和记录时间表，以确保采集和记录的及时性、合理性，极大提升了系统的整体运行效率水平，并且能够提供全方位的监听服务支持，系统会在信号异常的情况下自动发出警报，避免由于信号故障等问题所导致的音频丢失，还可以根据具体播出时间来调整录音时间；（4）视频信号采集站。在实际的应用过程中能够实现8路电视信号的采集和记录。操作中，利用视频信号采集站完整复合电视信号的音频信号码。一般来说，可以任意选择压缩码，并结合频道播出时间独立设计每个频道的采录时间。在具体的采录过程中，通常以MPEG压缩格式进行。页面可以呈现出所有视频图像且监听视频信号；同时，对各路信号的饱和度、色度及对比度进行整体调节，以确保采录质量[9]。

结语

总而言之，对于广播电视监测监管人员而言，无线电频率管理和监测监管工作至关重要，随着无线电技术得到普及，无线电频率资源日趋紧张，广播电视监测监管工作也同样开始朝着全自动化方向发展，这就需要有效运用语音识别技术，提升广播电视监测的工作质量及效率。因此，为能够确保广播电视监测监管事业的蓬勃发展，必须注重并加强语音识别技术在广播电视监测中有效、合理及大范围的应用。媒

参考文献

[1] 张君 . 语音识别技术在内容监管中的应用分析 [J]. 电视指南，2018（06）：254.

[2] 李志远 . 语音识别技术概述 [J]. 中国新通信，2018，20（17）：74–75.

[3] 刘刚，刘洋 . 语音识别技术在广播电台的应用构想 [J]. 中国有线电视，2015（11）：1300–1301.

[4] 刘林，孙成江，王正兴 . 语音识别技术在智能消防建设中的应用探索 [J]. 中国管理信息化，2018，21（23）：174–176.

[5]KJ.1029. 微软计算机语音识别技术开发取得重大突破 [J]. 军民两用技术与产品，2016（21）：20.

[6] 孙晓杰 . 嵌入式语音识别技术实现 [J]. 信息记录材料，2018，19（08）：118–119.

[7] 李粤 . 语音识别技术和语音合成技术对广播音频的冲击和应对 [J]. 传播力研究，2018，2（05）：133–136.

[8] 王海涛 . 广播电视监播系统中音频及语音数据处理技术研究 [D]. 西北工业大学，2007.

[9] 郭利刚，方土富 . 智能声音识别技术在广播电视广告监测中的应用 [J]. 广播与电视技术，2006（12）：72–74.

（作者单位：国家广播电视总局二八一台）